

ACÚSTICA DEL TRACTO VOCAL

Federico Miyara

1. INTRODUCCIÓN

El tracto vocal puede considerarse en primera aproximación como un conducto cuya sección transversal cambia con la posición a lo largo de su eje. Debido a los diversos mecanismos articulatorios ya comentados, la forma y dimensiones de dicha sección también experimentan, en principio, una evolución en el tiempo. No obstante, desde el punto de vista de los fenómenos acústicos involucrados esta evolución es muy lenta y entonces podemos suponer que el sistema es estacionario.

Nos proponemos estudiar el comportamiento del sonido en el interior de este tipo de conducto. Para asentar los conceptos físicos a utilizar, consideraremos inicialmente el caso más simple de un tubo cilíndrico.

2. LA ECUACIÓN DE ONDA

2.1. Ecuación de onda en un tubo de sección constante [1]

Consideremos un tubo cilíndrico de sección transversal A y extensión infinita como el indicado en dos instantes de tiempo (figura 1). Las dos regiones sombreadas corresponden a la misma porción de gas. En la parte superior de la figura la presión total es constante e igual a la presión atmosférica, P_0 . En la parte inferior la porción de aire

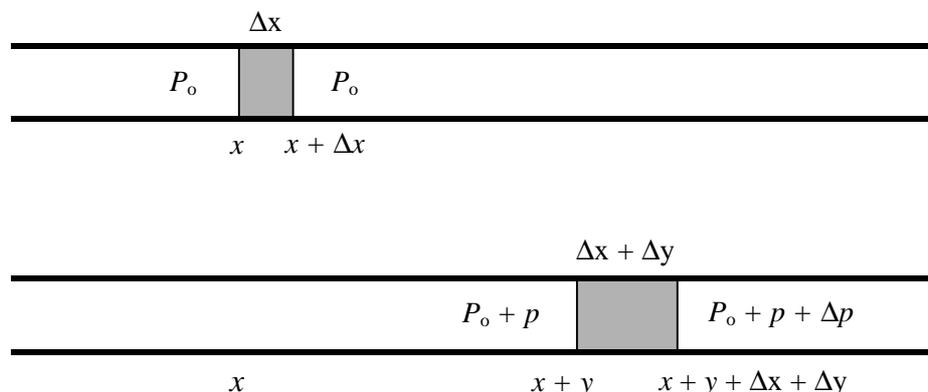


Figura 1. Un tubo de sección constante en dos instantes de tiempo diferentes. Arriba, en estado de equilibrio. Abajo, un pequeño volumen de aire se ha desplazado y expandido.

se ha desplazado una distancia y , y su espesor pasó de valer Δx a valer $\Delta x + \Delta y$. La presión se ha modificado en un valor p , en la cara izquierda y $p + \Delta p$ en la cara derecha, dando lugar a una fuerza neta que actúa sobre el volumen de gas. Tanto el desplazamiento como el incremento de presión son funciones de las dos variables x y t , es decir, $y(x, t)$ y $p(x, t)$. Llamaremos *presión sonora* al incremento de presión $p(x, t)$.

Podemos aplicar la segunda ley de Newton. Para ello tengamos en cuenta que si la densidad del aire es ρ_0 , entonces la masa de dicha porción es $\rho_0 A \Delta x$. La fuerza neta que actúa en la dirección del movimiento puede calcularse teniendo en cuenta que a la presión del lado izquierdo es $P_0 + p$, mientras que del lado derecho es $P_0 + p + \Delta p$:

$$(P_0 + p) A - (P_0 + p + \Delta p) A = -\Delta p A.$$

Entonces, dado que la aceleración es y_{tt} , resulta

$$-\Delta p A = \rho_0 A \Delta x \cdot y_{tt}.$$

Dividiendo por A y por Δx y pasando al límite cuando $\Delta x \rightarrow 0$, se obtiene

$$-p_x = \rho_0 y_{tt}. \quad (1)$$

Esta es una ecuación diferencial en derivadas parciales con dos funciones incógnitas: la presión p y el desplazamiento y . Se necesitará otra ecuación que las vincule para poder despejarlas. Dicha ecuación puede obtenerse mediante las relaciones que existen entre el volumen y la presión en un medio gaseoso. Si la temperatura fuera constante (proceso *isotérmico*), la relación sería la *ley de Boyle*:

$$P \cdot V = cte.$$

Pero el gradiente de temperatura es muy bajo y la conductividad térmica del aire también, por lo que el intercambio de calor es despreciable. El proceso resulta, por consiguiente, *adiabático*. Las compresiones adiabáticas responden a la ley

$$P \cdot V^\gamma = cte. \quad (2)$$

donde $\gamma = C_p/C_v$, es decir el cociente entre las capacidades caloríficas a presión constante y a volumen constante. Para gases diatómicos, como el aire, $\gamma \cong 1,4$. De la ecuación 2 podemos obtener la expresión incremental

$$\frac{dP}{P} = -\gamma \frac{dV}{V}. \quad (3)$$

En el presente caso, $dP = p$, $P = P_0$, $dV = A \cdot \Delta y$, $V = A \cdot \Delta x$, de donde

$$\frac{p}{P_0} = -\gamma \frac{\Delta y}{\Delta x}.$$

Pasando al límite cuando $\Delta x \rightarrow 0$, resulta,

$$\frac{P}{P_0} = -\gamma y_x. \quad (4)$$

Tenemos ahora dos ecuaciones, (1) y (4), que vinculan las incógnitas funcionales $p(x, t)$ e $y(x, t)$. En general estamos más interesados en la presión que en el desplazamiento, debido a que es la magnitud que puede medirse directamente con mayor facilidad (utilizando un micrófono como transductor). Derivando la ecuación (1) con respecto a x y la (4) dos veces con respecto a t , tenemos, finalmente,

$$p_{tt} = \frac{\gamma P_0}{\rho_0} p_{xx}, \quad (5)$$

o bien

$$p_{tt} = c^2 p_{xx}, \quad (6)$$

ecuación conocida como *ecuación de onda unidimensional*. La constante c es la *velocidad del sonido*.

Es posible demostrar que la solución general de la ecuación (6) está dada por

$$p(x, t) = f(x - c \cdot t) + g(x + c \cdot t). \quad (7)$$

donde f y g son dos funciones arbitrarias de una variable derivables dos veces.¹ Cada término de esta ecuación puede interpretarse de un modo muy sencillo. Por ejemplo, si en el término $f(x - c \cdot t)$ tomamos $t = 0$, tendremos la distribución inicial de presión, como se indica en la parte superior de la figura 2. Luego de un tiempo $t > 0$ la nueva distribución de presión es la que se indica en la parte inferior. Vemos que cualquier rasgo distintivo (por ejemplo un pico) de la presión se ha movido una distancia $c \cdot t$ hacia la derecha. Es importante notar que lo que se propaga es la distribución de presión, no la materia.

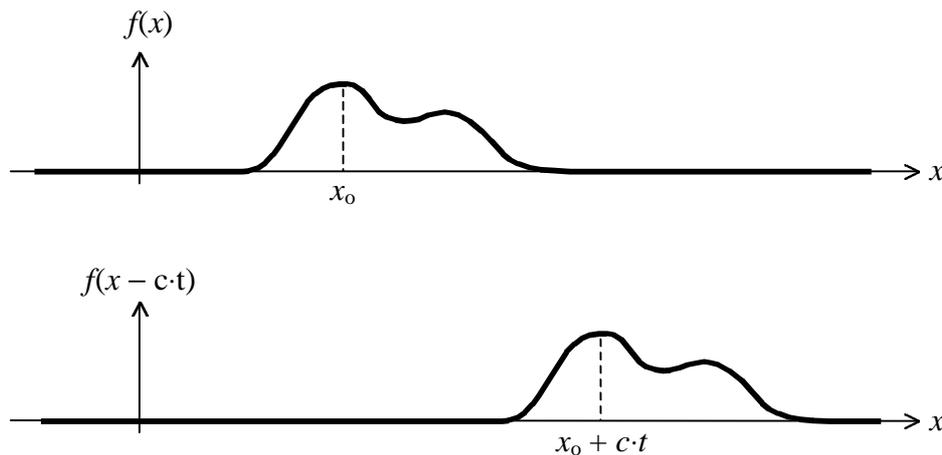


Figura 2. Propagación de una onda a lo largo del tubo. Arriba se indica la distribución inicial ($t = 0$) de la presión a lo largo del tubo. Abajo se observa cómo luego de un tiempo $t > 0$ la distribución de presión se ha desplazado una distancia $c \cdot t$.

¹ Es fácil verificar que la función dada por la ecuación (7) es solución, para lo cual basta calcular las derivadas correspondientes. Pero además, *todas* las soluciones de la ecuación (6) son de esta forma.

2.2. Ecuación de onda en un tubo de sección variable [2]

Consideremos ahora un tubo cuya sección transversal es $A(x)$ como el indicado en dos instantes de tiempo en la figura 3. Nuevamente, las dos regiones sombreadas corresponden a la misma porción de gas. En la parte superior de la figura se representa la situación de equilibrio, es decir en ausencia de perturbaciones. La presión total es constante e igual a la presión atmosférica, P_0 . En la parte inferior la porción de aire se ha desplazado una distancia y y se ha descomprimido desde un espesor Δx a uno $\Delta x + \Delta y$. Mientras tanto, la presión sufrió un incremento p en la cara izquierda y $p + \Delta p$ en la cara derecha, y la sección cambió de $A(x)$ a $A(x + y)$.²

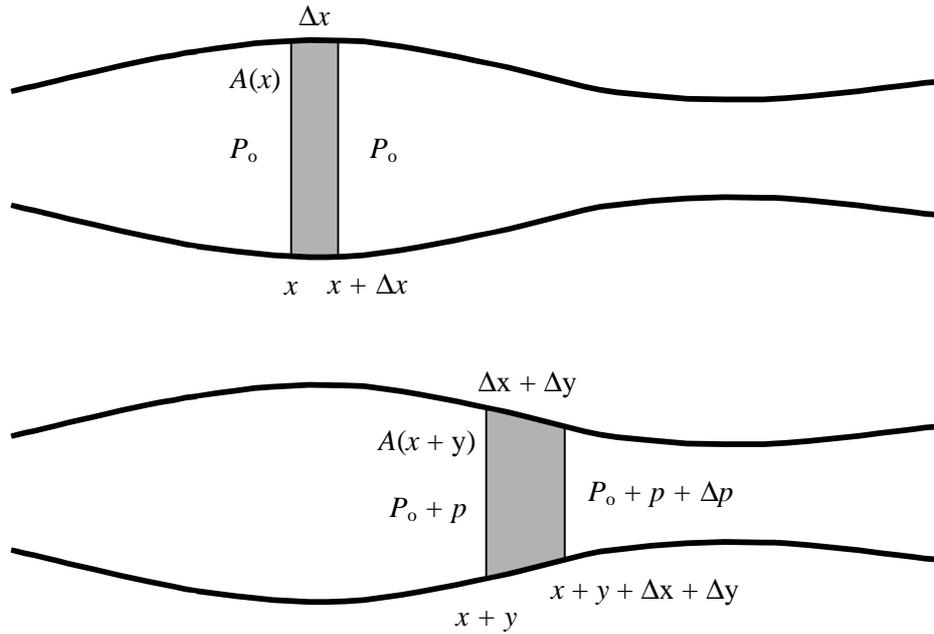


Figura 3. Un tubo de sección variable en dos instantes de tiempo diferentes. Arriba, en estado de equilibrio. Abajo, un pequeño volumen de aire se ha desplazado y expandido.

Aplicemos ahora la segunda ley de Newton a la porción de aire. En este caso la masa de es $\rho_0 A(x) \Delta x$. La fuerza neta que actúa en la dirección del movimiento puede calcularse a partir de las presiones del lado izquierdo y derecho:

$$(P_0 + p) A(x + y) - (P_0 + p + \Delta p) A(x + y) = -\Delta p A(x + y).$$

Dado que normalmente y es pequeño, puede aproximarse $A(x + y)$ mediante un desarrollo de Taylor de primer grado:

$$A(x + y) \cong A(x) + A'(x) y. \quad (8)$$

² A diferencia del caso anterior, en el que la sección era constante, es decir $A(x) = A$, ahora las secciones a uno y otro lado de la porción analizada son diferentes (por ejemplo, $A(x)$ y $A(x + \Delta x)$ en la situación de equilibrio). Sin embargo, dado que haremos tender Δx a 0, la diferencia no será importante. En cambio sí es importante diferenciar $A(x)$ y $A(x + y)$, ya que y , pese a ser pequeño, no tenderá a 0.

Entonces la fuerza neta actuante vale

$$-\Delta p A(x+y) \cong -\Delta p [A(x) + A'(x)y].$$

Aplicando la ley de Newton,

$$-\Delta p A(x+y) = \rho_o A(x) \Delta x \cdot y_{tt}.$$

Dividiendo por $A(x)$ y por Δx y pasando al límite cuando $\Delta x \rightarrow 0$, se obtiene

$$-p_x \left[1 + \frac{A'(x)}{A(x)} y \right] = \rho_o y_{tt}. \quad (9)$$

Esta ecuación diferencial en derivadas parciales es no lineal, lo cual complica notablemente su resolución. Si y es pequeño (lo cual sucede para sonidos no muy intensos) el incremento relativo del área entre x y $x+y$ es despreciable y entonces se puede aproximar por

$$-p_x = \rho_o y_{tt}. \quad (10)$$

Ahora debemos particularizar a este nuevo caso la ecuación de la compresión adiabática (ecuación (3)). Igual que antes, $dP = p$, $P = P_o$. El volumen inicial será ahora $V(x) = A(x) \cdot \Delta x$ y el incremento de volumen

$$dV = A(x+y) \cdot (\Delta x + \Delta y) - A(x) \cdot \Delta x$$

$$dV \cong [A(x) + A'(x)y] \cdot (\Delta x + \Delta y) - A(x) \cdot \Delta x.$$

Reemplazando

$$\frac{p}{P_o} = -\frac{\gamma}{A(x)} \frac{[A(x) + A'(x)y] \Delta y + A'(x)y \Delta x}{\Delta x}.$$

Pasando al límite cuando $\Delta x \rightarrow 0$, resulta,

$$\frac{p}{P_o} = -\frac{\gamma}{A(x)} [[A(x) + A'(x)y] y_x + A'(x)y].$$

Aquí podemos efectuar la misma aproximación anterior, es decir $A'(x)y \ll A(x)$, debido al pequeño incremento relativo de área causado por un desplazamiento tan pequeño como lo es y para intensidades moderadas. Resulta, finalmente

$$\frac{p}{P_o} = -\frac{\gamma}{A(x)} [A(x) y_x + A'(x)y],$$

es decir

$$\frac{p}{P_o} = -\frac{\gamma}{A(x)} [A(x) y]_x. \quad (11)$$

Como estamos interesados en una ecuación diferencial con la presión $p(x, t)$ como única incógnita, debemos ahora eliminar la incógnita y de las ecuaciones (10) y (11). Para ello derivamos la (11) dos veces con respecto a t . Luego invertimos el orden de derivación, e introducimos las derivadas temporales dentro del corchete. Como $A(x)$ es independiente del tiempo, dentro del corchete aparece y_{tt} , que puede ser reemplazado por el valor obtenido de la ecuación (10). Resulta:

$$p_{tt} = \frac{c^2}{A(x)} [A(x) p_x]_x. \quad (12)$$

donde $c^2 = \gamma P_0 / \rho_0$ es, igual que antes, la velocidad del sonido. Esta ecuación se conoce como *ecuación de la corneta*, o *ecuación de la bocina*. En el caso en que $A(x) = A$, es fácil verificar que se reduce al caso del tubo cilíndrico.

A modo de ejemplo resolveremos el caso de una *bocina exponencial* como las utilizadas en los parlantes de compresión (compression drivers). Supondremos entonces que

$$A(x) = A_0 e^{x/L}, \quad (13)$$

donde A_0 es el área de la sección en el origen y L la longitud para la cual el área ha crecido e veces. Reemplazando en la ecuación 12 y operando,

$$\frac{1}{c^2} p_{tt} = p_{xx} + \frac{1}{L} p_x. \quad (14)$$

Por la forma especial de $A(x)$ la ecuación quedó reducida a una ecuación lineal a coeficientes constantes. Cuando $L \rightarrow \infty$ la ecuación se aproxima a la ecuación de onda para el tubo cilíndrico.

De especial interés en este tipo de ecuaciones es la obtención de *soluciones estacionarias*, es decir, soluciones en las que las variables se encuentran separadas en dos factores multiplicativos:

$$p(x, t) = X(x) \cdot T(t). \quad (15)$$

Las soluciones estacionarias forman una *base de soluciones*, es decir, un conjunto de soluciones tal que cualquier otra solución puede expresarse como combinación lineal de ellas. Además, constituyen los *modos normales de oscilación* característicos del sistema. En este caso se cumple

$$\frac{1}{c^2} X T'' = X'' T + \frac{1}{L} X' T.$$

Dividiendo por $X T$ se puede expresar como

$$\frac{1}{c^2} \frac{T''}{T} = \frac{X''}{X} + \frac{1}{L} \frac{X'}{X}.$$

Dado que el lado izquierdo es una función sólo de t y el derecho sólo de x , ambos deben ser iguales a una misma constante que llamamos $-\lambda^2$:

$$\frac{1}{c^2} \frac{T''}{T} = \frac{X'' + \frac{1}{L} X'}{X} = -\lambda^2.$$

Se obtiene el siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{cases} T'' + \lambda^2 c^2 T = 0 \\ X'' + \frac{1}{L} X' + \lambda^2 X = 0 \end{cases} \quad (16)$$

Su solución puede llevarse a cabo por cualquiera de los métodos conocidos, obteniéndose

$$T(t) = T_o \text{sen}(Ic t + \phi_o) \quad (17)$$

$$X(x) = X_o e^{-x/2L} \text{sen}\left(\frac{x}{2L} \sqrt{4\lambda^2 L^2 - 1} + \psi_o\right). \quad (18)$$

Consideremos el caso particular de una bocina exponencial de longitud l , como se muestra en la figura 4. Si suponemos que la bocina está abierta en sus extremos, la presión será muy pequeña en ellos, pudiendo considerarse nula. Esto impone condiciones de contorno dadas por

$$p(0, t) = p(l, t) = 0. \quad (19)$$

para todo tiempo t . Esto significa que

$$\Psi_o = 0, \quad (20)$$

$$\frac{l}{2L} \sqrt{4\lambda^2 L^2 - 1} = k\pi. \quad (21)$$

De esta ecuación se pueden obtener los valores admisibles de λ :

$$\lambda = \frac{1}{2L} \sqrt{1 + \left(\frac{2k\pi L}{l}\right)^2}, \quad (22)$$

donde $k = 1, 2, \dots$. La presión sonora viene dada, entonces, por

$$p(x, t) = P e^{-x/2L} \text{sen}\left(\frac{k\pi}{l} x\right) \text{sen}\left(\frac{c}{2L} \sqrt{1 + \left(\frac{2k\pi L}{l}\right)^2} t + \phi_o\right). \quad (23)$$

En la parte inferior de la figura 4 se indica la distribución de la presión sonora a lo largo de la bocina para tres valores de k para un valor de t que hace máximo el factor dependiente del tiempo. Las frecuencias no están distribuidas armónicamente, salvo en el caso en que $L \gg l$, es decir, cuando la bocina tiende a ser cilíndrica.

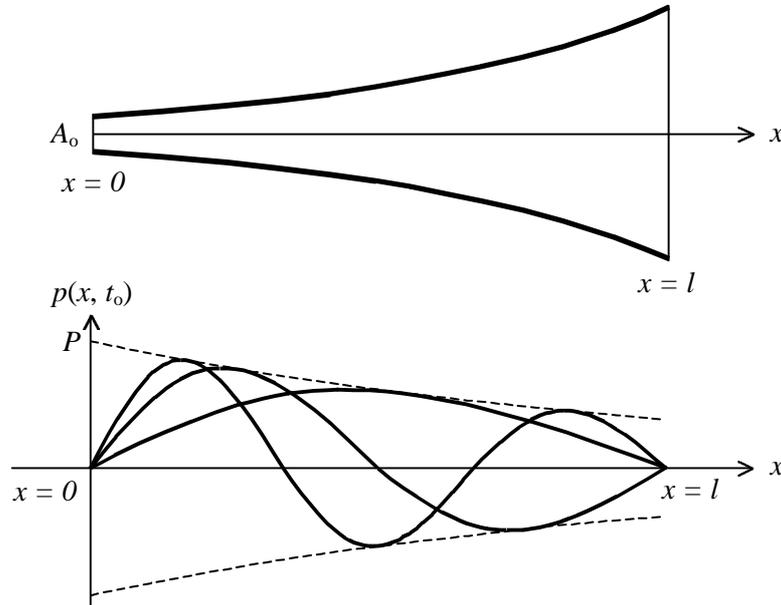


Figura 4. Arriba, una bocina exponencial. Abajo, distribución de la presión sonora en un instante para tres soluciones estacionarias con $n = 1, 2$ y 3 .

2.2. El tracto vocal con pérdidas [3]

El análisis realizado anteriormente supuso un tubo idealizado carente de todo tipo de pérdidas. Un conducto real exhibe al menos dos fenómenos que obligan a modificar las ecuaciones. El primero es la fricción viscosa; el segundo, la absorción en las paredes.

En lugar de efectuar el análisis en términos del desplazamiento $y(x, t)$ es conveniente hacerlo en términos de la velocidad de las partículas $v(x, t)$ y de la *velocidad volumétrica*, o *caudal*, $u(x, t)$. Notemos que $v(x, t)$ puede interpretarse de dos maneras: como la velocidad en el instante t de la partícula que *en equilibrio* estaría en x , o como la velocidad de la partícula que en el instante t se encuentra *efectivamente* en x . De acuerdo con la notación que veníamos utilizando, la primera interpretación conduce a

$$v(x, t) = \dot{y}_i(x, t), \quad (24)$$

mientras la segunda, a

$$v(x, t) = \dot{y}_i(x, t) - \dot{y}_{ix}(x, t) y(x, t), \quad (25)$$

Sin embargo, el término agregado en la ecuación (25) es despreciable, por lo cual utilizaremos la segunda interpretación pero aceptaremos como válida la ecuación (24).

La velocidad volumétrica o caudal, $u(x, t)$ es el volumen que cruza una sección de área $A(x)$ por unidad de tiempo. Suponiendo que la velocidad es normal a la sección (lo cual sucede pues suponemos que las ondas son planas), $u(x, t)$ está dada por

$$u(x, t) = v(x, t) A(x). \quad (26)$$

Entonces las ecuaciones (10) y (11) adoptan la forma simétrica siguiente:

$$p_x = -\frac{\rho_0}{A(x)} u_t \quad (27)$$

$$p_t = -\frac{\gamma P_0}{A(x)} u_x. \quad (28)$$

La modificación de estas ecuaciones para tener en cuenta la fricción y la absorción en las paredes se simplifica mucho si en lugar de continuar trabajando en el dominio tiempo lo hacemos en el dominio frecuencia, utilizando para ello la transformación de Laplace. Tengamos en cuenta que la transformación de Laplace clásica se aplica a funciones de una sola variable, en tanto que aquí tenemos funciones de dos variables. Si consideramos la variable espacial como un parámetro fijo en el momento de transformar, obtendremos una transformada dependiente de ese mismo parámetro. Así, la transformada de $p(x, t)$ será $P(x, s)$, y la transformada de $u(x, t)$ resultará $U(x, s)$. Entonces

$$P_x = -\frac{\rho_0 s}{A(x)} U \quad (29)$$

$$sP = -\frac{\gamma P_0}{A(x)} U_x, \quad (30)$$

donde las derivadas temporales quedaron reemplazadas por el factor s .

Tengamos en cuenta que $A \cdot P_x dx$ representa la fuerza neta sobre una capa de aire de espesor dx debida al cambio de presión, en tanto que $\rho_0 s U$ es la fuerza de aceleración que se opone a la anterior. En caso de haber fricción viscosa, se agrega una nueva fuerza que tiende a contrarrestar a la fuerza de presión, que resulta proporcional a la velocidad. El factor de proporcionalidad puede depender de la frecuencia y de la posición, lo que hubiera sido difícil de tener en cuenta en el dominio tiempo. Por ejemplo, es de suponer que en un estrangulamiento la fricción sea mucho más alta que en una cavidad de gran sección. La ecuación 29 puede escribirse, entonces, como

$$P_x = -\frac{\rho_0 s}{A(x)} U - R(x, s) U \quad (31)$$

Para tener en cuenta las pérdidas por absorción en las paredes, reescribamos la ecuación 30 de la forma siguiente:

$$U_x = -\frac{A(x)}{\gamma P_0} sP. \quad (32)$$

Observemos que $U_x dx$ representa la variación de caudal entre la cara derecha y la izquierda de una capa de aire de espesor dx . La ecuación 30 establece uno de los mecanismos por los que se pierde caudal: el hecho de que la presión está aumentando, lo cual hace disminuir el volumen de una dada masa de aire. El otro mecanismo, que pretendemos incorporar, es el aumento de la sección total debido a la presión sobre las paredes del tracto cuando éstas no son rígidas. Este fenómeno se ilustra en la figura 5.

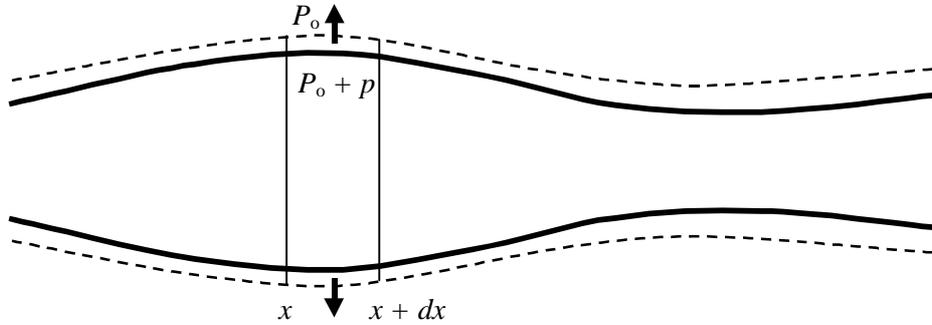


Figura 5. Efecto de la presión sobre las paredes no rígidas del tracto. Se ha supuesto que los diferentes puntos reaccionan localmente.

Para simplificar el análisis deberemos suponer que las paredes interiores del tracto responden localmente a la presión, es decir, lo hacen con independencia de lo que suceda con un punto vecino. Si bien en realidad esto no sucede, el hecho de que por continuidad la presión no varíe bruscamente de un punto a otro vecino lo compensa.

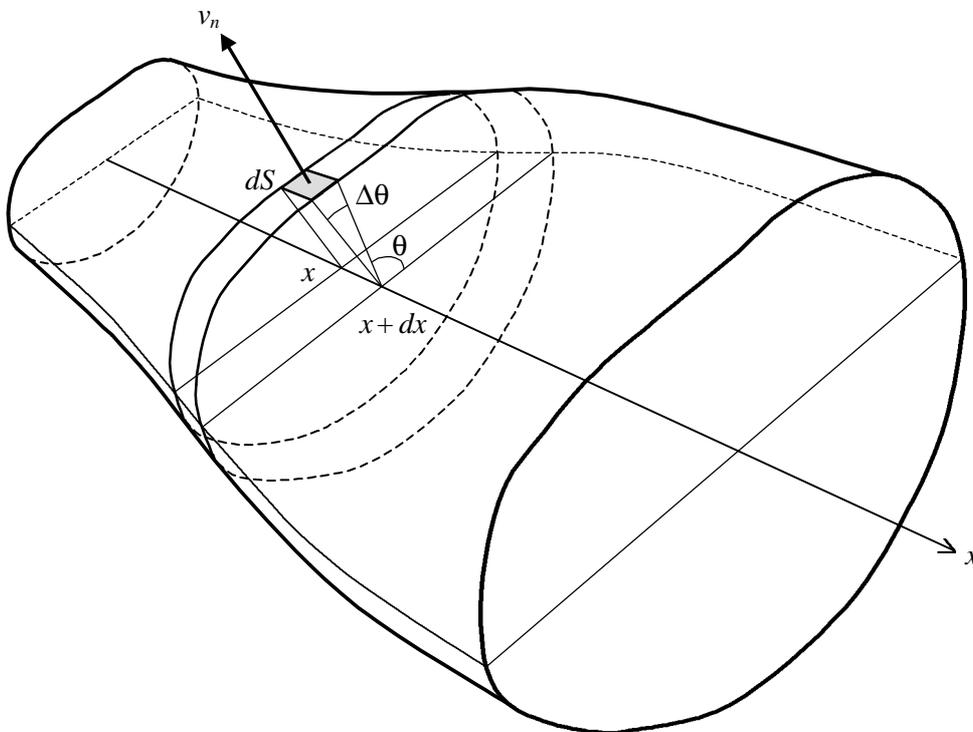


Figura 6. Geometría para el cálculo de la admitancia acústica total de la pared del tracto.

La reacción local de una superficie ante la presión se traduce en el concepto de *impedancia acústica* como cociente entre las transformadas de Laplace de la presión y la velocidad, y en el concepto dual de *admitancia acústica*, cociente entre las transformadas de la velocidad y la presión. Con referencia a la figura 6, si la presión sobre la pared en un punto ubicado en coordenadas cilíndricas (θ, x) es $P_p(\theta, x, s)$ y la velocidad de desplazamiento de la pared normal a la superficie es $V_n(\theta, x, s)$, entonces la admitancia puntual es

$$Y^*(\theta, x, s) = \frac{V_n(\theta, x, s)}{P_p(\theta, x, s)}. \quad (33)$$

Dado que la longitud de onda es mucho mayor que las dimensiones transversales del tracto vocal,³ podemos suponer que el campo sonoro es plano, por lo cual la presión sonora es aproximadamente constante en toda la sección, es decir $P_p(\theta, x, s) \cong P(x, s)$, de manera que

$$V_n(\theta, x, s) = Y^*(\theta, x, s) P(x, s). \quad (34)$$

Nos proponemos calcular el caudal que escapa por la pared a causa del corrimiento de ésta. Consideremos para ello el caudal que escapa por el contorno de la delgada lámina de espesor dx de la figura 6. Podemos escribir el elemento de área como

$$dS = S(\theta, x) d\theta dx,$$

donde $S(\theta, x)$ depende de la forma del conducto. El caudal que escapa por dicho contorno será

$$\begin{aligned} dU_{pared} &= dx \int_0^{2\pi} V_n(\theta, x, s) S(\theta, x) d\theta = \\ &= dx \int_0^{2\pi} Y^*(\theta, x, s) P(x, s) S(\theta, x) d\theta = \\ &= P(x, s) dx \int_0^{2\pi} Y^*(\theta, x, s) S(\theta, x) d\theta \end{aligned}$$

Llamando *admitancia total* en el punto x a

$$Y(x, s) = \int_0^{2\pi} Y^*(\theta, x, s) S(\theta, x) d\theta, \quad (35)$$

y dividiendo por dx podemos escribir

$$U_{pared,x}(x, s) = Y(x, s) P(x, s). \quad (36)$$

Este valor representa las pérdidas de caudal debidas a que la presión en aumento está empujando las paredes perimetrales de la lámina hacia afuera, y debe restarse en la ecuación (30):

³ Las frecuencias de interés para la voz humana son inferiores a 4 kHz, la longitud de onda resultante es mayor de 8,6 cm, mientras que las dimensiones del tracto vocal no superan en general los 2 ó 3 cm.

$$U_x = -\frac{A(x)}{\gamma P_0} sP - Y(x, s) \cdot P. \quad (37)$$

Hemos así incorporado, en las ecuaciones (29) y (37), los efectos de la fricción y la absorción en las paredes internas del tracto. Combinándolas obtenemos una ecuación que involucra únicamente la presión $P(x, s)$.

$$\left(\frac{A}{\rho_0 s + A \cdot R} P_x \right)_x = \left(\frac{A s}{\rho_0 c^2} + Y \right) P. \quad (38)$$

Si bien ésta es una ecuación en derivadas parciales, en la práctica se comporta como si fuera una ecuación ordinaria, ya que sólo aparecen las derivadas con respecto a una sola variable (la x). Las derivadas respecto a la variable t quedaron absorbidas al aplicar la transformación de Laplace. Es, entonces, una ecuación homogénea de segundo orden a coeficientes variables y dependientes además del parámetro s . La solución general es de la forma

$$P(x, s) = a g(x, s) + b h(x, s), \quad (39)$$

donde g y h son dos soluciones particulares linealmente independientes, y, por aplicación de la ecuación 31,

$$U(x, s) = \frac{-A}{\rho_0 s + A \cdot R} (a g_x(x, s) + b h_x(x, s)). \quad (40)$$

3. REPRESENTACIÓN MATRICIAL

Es interesante analizar el tracto vocal como una caja negra con una entrada (la glotis) y una salida (los labios). Dado que existen dos variables de entrada y dos de salida (las respectivas presiones y velocidades de volumen), el problema es enteramente similar al de un cuadripolo eléctrico, y su descripción puede realizarse por medio de una representación matricial [2].

3.1. Caso general

Si suponemos que la entrada está en $x = 0$ y la salida en $x = L$, tendremos

$$P_{ent} = P(0, s) = a g(0, s) + b h(0, s), \quad (41)$$

$$U_{ent} = U(0, s) = \frac{-A}{\rho_0 s + A \cdot R} (a g_x(0, s) + b h_x(0, s))$$

$$P_{sal} = P(L, s) = a g(L, s) + b h(L, s). \quad (42)$$

$$U_{sal} = U(L, s) = \frac{-A}{\rho_0 s + A \cdot R} (a g_x(L, s) + b h_x(L, s))$$

Tenemos 4 ecuaciones y 6 incógnitas (a , b , P_{sal} , U_{sal} , P_{ent} y U_{ent}). Podemos elegir dos de ellas como parámetros independientes, y luego obtener los restantes. Suponiendo conocidos P_{sal} y U_{sal} podemos despejar las constantes a y b de las ecuaciones (42), y de las (41) obtener P_{ent} y U_{ent} . Resulta, aplicando la regla de Cramer,

$$a = \frac{\begin{vmatrix} P_{sal} & h(L, s) \\ -\frac{\rho_0 s + AR}{A} U_{sal} & h_x(L, s) \end{vmatrix}}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}}$$

$$b = \frac{\begin{vmatrix} g(L, s) & P_{sal} \\ g_x(L, s) & -\frac{\rho_0 s + AR}{A} U_{sal} \end{vmatrix}}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}}. \quad (43)$$

Entonces

$$\begin{pmatrix} P_{ent} \\ U_{ent} \end{pmatrix} = \frac{\begin{pmatrix} g(0, s) & h(0, s) \\ \frac{-A}{\rho_0 s + AR} g_x(0, s) & \frac{-A}{\rho_0 s + AR} h_x(0, s) \end{pmatrix}}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}} \begin{pmatrix} \begin{vmatrix} P_{sal} & h(L, s) \\ -\frac{\rho_0 s + AR}{A} U_{sal} & h_x(L, s) \end{vmatrix} \\ \begin{vmatrix} g(L, s) & P_{sal} \\ g_x(L, s) & -\frac{\rho_0 s + AR}{A} U_{sal} \end{vmatrix} \end{pmatrix} \quad (44)$$

o bien

$$\begin{pmatrix} P_{ent} \\ U_{ent} \end{pmatrix} = \frac{\begin{pmatrix} g(0, s) & h(0, s) \\ \frac{-A}{\rho_0 s + AR} g_x(0, s) & \frac{-A}{\rho_0 s + AR} h_x(0, s) \end{pmatrix}}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}} \begin{pmatrix} h_x(L, s) P_{sal} + h(L, s) \frac{\rho_0 s + AR}{A} U_{sal} \\ -g_x(L, s) P_{sal} - g(L, s) \frac{\rho_0 s + AR}{A} U_{sal} \end{pmatrix}$$

Esto puede expresarse en la forma

$$\begin{pmatrix} P_{ent} \\ U_{ent} \end{pmatrix} = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \begin{pmatrix} P_{sal} \\ U_{sal} \end{pmatrix} = \mathbf{K} \begin{pmatrix} P_{sal} \\ U_{sal} \end{pmatrix}, \quad (45)$$

donde los parámetros K_{ij} son análogos a los parámetros A, B, C y D de las matrices de transmisión de los cuadripolos, utilizadas para describir conexiones en cascada (por ejemplo, en líneas de transmisión). Los valores resultantes son:

$$K_{11} = \frac{g(0, s)h_x(L, s) - h(0, s)g_x(L, s)}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}} \quad (46)$$

$$K_{12} = \frac{\rho_0 s + AR}{A} \frac{g(0, s)h(L, s) - h(0, s)g(L, s)}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}} \quad (47)$$

$$K_{21} = \frac{-A}{\rho_0 s + AR} \frac{g_x(0, s)h_x(L, s) - h_x(0, s)g_x(L, s)}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}} \quad (48)$$

$$K_{22} = -\frac{g_x(0, s)h(L, s) - h_x(0, s)g(L, s)}{\begin{vmatrix} g(L, s) & h(L, s) \\ g_x(L, s) & h_x(L, s) \end{vmatrix}}. \quad (49)$$

Los coeficientes K_{ij} son funciones de s , y en la generalidad de los casos son demasiado complejos como para su tratamiento analítico mediante fórmulas cerradas. Si bien en el mejor de los casos las funciones $A(x, s)$, $R(x, s)$ e $Y(x, s)$ podrían determinarse experimentalmente y luego aproximarse mediante sendas expresiones analíticas, en la práctica las dificultades para medir en forma directa la resistencia viscosa y la admitancia de las paredes del tracto vocal son enormes.

En la próxima sección analizaremos un caso sencillo al cual podremos reducir los casos reales.

3.2. Caso de un tubo uniforme

Supondremos ahora el caso particular de un tubo de sección, fricción y admitancia constantes a lo largo de toda su longitud para cada valor de s , es decir

$$A(x, s) = A_0(s) \quad (50)$$

$$R(x, s) = R_0(s) \quad (51a)$$

$$Y(x, s) = Y_0(s) \quad (51b)$$

Entonces la ecuación (38) puede expresarse como

$$P_{xx} = \sigma^2 P, \quad (52)$$

donde

$$\sigma^2(s) = (\rho_0 s + AR) \left(\frac{s}{\rho_0 c^2} + \frac{Y}{A} \right). \quad (53)$$

La ecuación 52 admite una sencilla solución general de la forma

$$P(x, s) = a Ch(\sigma x) + b Sh(\sigma x). \quad (54)$$

Con las funciones $g(x, s) = Ch(\sigma x)$ y $h(x, s) = Sh(\sigma x)$ se pueden calcular los coeficientes K_{ij} . Resulta

$$\mathbf{K} = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \begin{pmatrix} Ch(\sigma L) & \frac{\rho_0 s + AR}{A\sigma} Sh(\sigma L) \\ \frac{A\sigma}{\rho_0 s + AR} Sh(\sigma L) & Ch(\sigma L) \end{pmatrix}. \quad (55)$$

En este caso sencillo fue posible obtener una expresión cerrada para la matriz de transmisión, siempre y cuando se conozcan A , R e Y . Aun en el caso ideal en que $R = 0$ e $Y = 0$ (resultando $\sigma = s/c$) los coeficientes obtenidos no son funciones racionales. Esto es característico de las líneas de transmisión a parámetros distribuidos. De hecho, el tubo rígido sin pérdidas se comporta como una línea de transmisión eléctrica puramente reactiva (con el análogo a la impedancia característica igual a la impedancia acústica del aire, $\rho_0 c$, dividida por la sección A para convertir la velocidad lineal en volumétrica).

3.3. Descomposición del tracto vocal en tubos uniformes

A pesar de que el tracto vocal no es uniforme, puede analizarse de una manera simplificada si se lo reduce a un cierto número de componentes uniformes conectadas una a continuación de la otra, como se muestra en la figura 7.

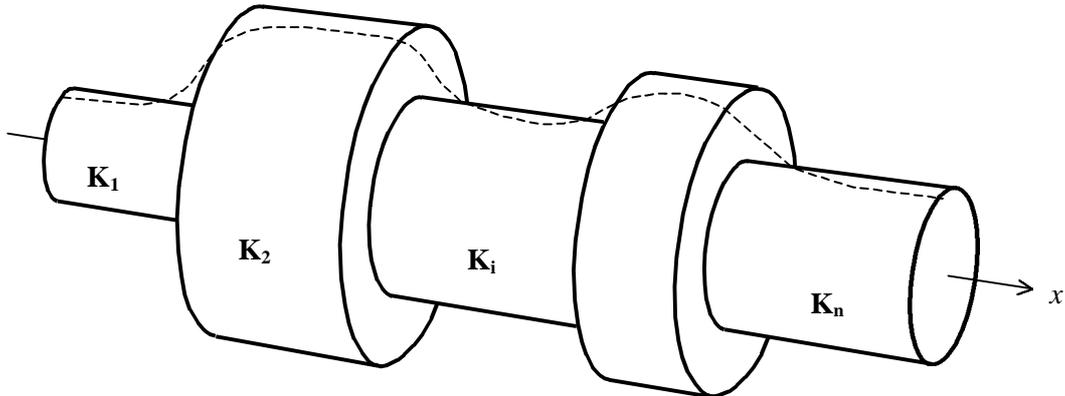


Figura 7. Aproximación del tracto vocal mediante una interconexión de tubos de propiedades independientes de la posición, cada uno con su matriz de transmisión \mathbf{K}_i .

Si llamamos $\mathbf{K}_1, \dots, \mathbf{K}_n$ a las matrices de transmisión de cada segmento, podemos obtener la matriz total como el producto matricial de todas ellas.

$$\mathbf{K} = \mathbf{K}_1 \cdot \mathbf{K}_2 \cdot \dots \cdot \mathbf{K}_n. \quad (56)$$

Esto simplifica notablemente el análisis en los casos en que sea posible descomponer un conducto en secciones aproximadamente constantes. La ecuación (56) sigue cumpliéndose aún cuando las secciones sean variables.

3.4. Efecto de la impedancia de radiación

Al igual que en el caso de una línea de transmisión eléctrica, es interesante el efecto de carga de la impedancia de radiación. Para baja frecuencia (debajo de 800 Hz) esta impedancia vale aproximadamente [3]:

$$Z_R(s) = \frac{P_{labios}(s)}{U_{labios}(s)} \cong C \cdot s, \quad (57)$$

relación que representa la menor eficiencia en baja frecuencia de un emisor pequeño.

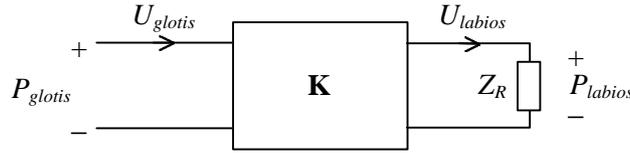


Figura 8. Un tracto vocal, representado como cuadripolo acústico, cargado por la impedancia de radiación.

Podemos escribir la ecuación (45) en la forma

$$\begin{pmatrix} P_{ent} \\ U_{ent} \end{pmatrix} = \mathbf{K} \begin{pmatrix} P_{sal} \\ U_{sal} \end{pmatrix} = \mathbf{K} \begin{pmatrix} Z_R \\ 1 \end{pmatrix} U_{sal}, \quad (58)$$

de donde

$$P_{glotis} = (K_{11}Z_R + K_{12}) U_{labios} \quad (59)$$

$$U_{glotis} = (K_{21}Z_R + K_{22}) U_{labios} \quad (60)$$

es decir

$$\frac{P_{glotis}}{U_{labios}} = K_{11}Z_R + K_{12} \quad (61)$$

$$\frac{U_{glotis}}{U_{labios}} = K_{21}Z_R + K_{22}. \quad (62)$$

Invirtiendo esta última, se la puede representar como función de transferencia:

$$\frac{U_{labios}}{U_{glotis}} = \frac{1}{K_{21}Z_R + K_{22}}. \quad (63)$$

Los ceros σ_i del denominador son los *formantes* del tracto. Si los ceros son complejos, es decir $\sigma_i = \alpha_i + j\omega_i$, entonces ω_i son las frecuencias (angulares) para las que la transferencia tiene máximos locales o resonancias (figura 9). Para cada ω_i , la correspondiente parte real α_i es (en valor absoluto) la mitad del ancho de banda (medido entre los puntos de -3 dB) de la campana de resonancia.

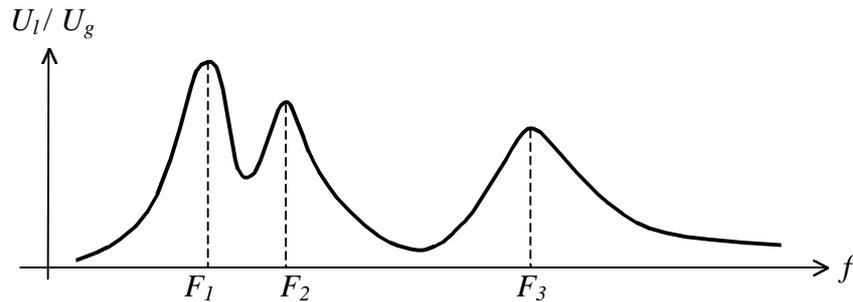


Figura 9. Formantes del tracto vocal para una determinada posición de los órganos articulatorios.

4. MODELO DE PREDICCIÓN LINEAL

Los modelos anteriores permiten una representación analítica de los fenómenos acústicos involucrados en la emisión vocal, pero no son fácilmente aprovechables para el tratamiento de la mayoría de los problemas prácticos, ya que no es sencillo obtener las funciones involucradas. Particularmente deseable sería un modelo apto para su tratamiento digital y que tuviera pocos parámetros de fácil obtención a partir de un análisis de ejemplos de emisiones. Un modelo de este tipo es el modelo de predicción lineal.

Supongamos primero el caso del tracto vocal sin pérdidas, es decir que $R = 0$ e $Y = 0$. En este caso, de la ecuación (53) se obtiene

$$\sigma = \frac{s}{c} \quad (64)$$

y entonces

$$\mathbf{K}_i = \begin{pmatrix} Ch\left(\frac{L_i s}{c}\right) & \frac{\rho_0 c}{A_i} Sh\left(\frac{L_i s}{c}\right) \\ \frac{A_i}{\rho_0 c} Sh\left(\frac{L_i s}{c}\right) & Ch\left(\frac{L_i s}{c}\right) \end{pmatrix}. \quad (65)$$

Podemos suponer que subdividimos el tracto vocal en N segmentos todos de igual longitud, de modo que $L_i = L/N$. Esto no implica pérdida de generalidad, ya que puede adaptarse un tamaño que sea suficientemente pequeño como para representar con sufi-

ciente aproximación todas las variaciones importantes en la sección del tracto. En ese caso

$$\mathbf{K}_i = \begin{pmatrix} Ch\left(\frac{Ls}{Nc}\right) & \frac{\rho_0 c}{A_i} Sh\left(\frac{Ls}{Nc}\right) \\ \frac{A_i}{\rho_0 c} Sh\left(\frac{Ls}{Nc}\right) & Ch\left(\frac{Ls}{Nc}\right) \end{pmatrix}. \quad (66)$$

Supondremos también que la impedancia de radiación también puede despreciarse y adoptaremos $Z_R = 0$. Entonces, de (63) resulta

$$\frac{U_{labios}}{U_{glotis}} = \frac{1}{K_{22}}, \quad (67)$$

de manera que para encontrar la función de transferencia entre el caudal en la glotis (entrada) y en los labios (salida) bastará hallar K_{22} . La matriz de transmisión de todo el tracto vocal, según la ecuación (56) puede expresarse como

$$\mathbf{K} = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \begin{pmatrix} K_{11,1} & K_{12,1} \\ K_{21,1} & K_{22,1} \end{pmatrix} \Lambda \begin{pmatrix} K_{11,N} & K_{12,N} \\ K_{21,N} & K_{22,N} \end{pmatrix}. \quad (68)$$

Se puede comprobar que K_{22} es un polinomio de 2^{N-1} términos, homogéneo en los $K_{hk,i}$, es decir en el que cada término tiene exactamente N factores $K_{hk,i}$:

$$K_{22} = \sum_{k=1}^{2^{N-1}} \left(\prod_{i=1}^N K_{a_k i b_{k i}, i} \right), \quad (69)$$

donde los índices $a_{k i}$ y $b_{k i}$ son iguales a 1 ó 2 y cumplen además que $a_{k1} = b_{kN} = 2$ y $a_{k(i+1)} = b_{k i}$ (por ejemplo, para $N = 4$, uno de los términos es $K_{21,1} \cdot K_{12,2} \cdot K_{22,3} \cdot K_{22,4}$)⁴. Por otra parte, cada elemento de las matrices de (68) depende de s en la forma de un seno o coseno hiperbólico. Dichas funciones pueden expresarse de la siguiente manera

$$Ch x = \frac{e^x + e^{-x}}{2} = \frac{1 + e^{-2x}}{2 e^{-x}}, \quad (70)$$

$$Sh x = \frac{e^x - e^{-x}}{2} = \frac{1 - e^{-2x}}{2 e^{-x}}, \quad (71)$$

lo cual permite expresar

$$K_{a_k i b_{k i}, i} = C_{k i} \frac{1 \pm e^{-2Ls / Nc}}{2e^{-Ls / Nc}} \quad (72)$$

⁴ En forma más general, K_{mh} requiere que $a_{k1} = m$ y que $b_{kN} = h$, siendo las otras condiciones las mismas.

donde la elección del signo + o el - depende de si el elemento contenía un *Sh* o un *Ch*.
 Resulta

$$\mathbf{K}_{22} = \frac{1}{e^{-Ls/c}} \sum_{k=1}^{2^{N-1}} C_k \prod_{i=1}^N \left(1 \pm e^{-2Ls/Nc}\right). \quad (73)$$

La suma es un polinomio de grado N en $e^{-2Ls/Nc}$, por lo cual

$$\mathbf{K}_{22} = \frac{1}{e^{-Ls/c}} \sum_{k=0}^N a_k e^{\frac{-2Ls}{Nc}k}. \quad (74)$$

Entonces

$$G_U(s) = \frac{U_{labios}}{U_{glotis}} = \frac{e^{-\frac{Ls}{c}}}{\sum_{k=0}^N a_k e^{\frac{-2Ls}{Nc}k}}. \quad (75)$$

Recordemos que la transformada de Laplace de un retardo temporal se obtiene multiplicando la transformada por una exponencial. Podemos poner esto de manifiesto del siguiente modo:

$$U_{labios}(s) \sum_{k=0}^N a_k e^{-\frac{2Ls}{Nc}k} = U_{glotis}(s) e^{-\frac{N}{2} \frac{2Ls}{Nc}}. \quad (76)$$

Llamando $T = 2L/Nc$ y antitransformando,

$$\sum_{k=0}^N a_k u_{labios}(t - kT) = u_{glotis}\left(t - \frac{N}{2}T\right). \quad (77)$$

Esta expresión indica que hay una relación directa entre el caudal en la glotis retardado un tiempo $\frac{1}{2}NT$ y el caudal en los labios en sus versiones retardadas unos tiempos kT con $k = 1, \dots, N$.

La ecuación (77) tiene el aspecto de una ecuación diferencia, salvo porque el tiempo t es aquí una variable continua y no discreta. Podemos transformarla en discreta si tomamos muestras a intervalos iguales a T , es decir con una tasa de muestreo $f_s = 1/T$. Para que esto sea factible sin introducir distorsiones por aparición de frecuencias alias debe elegirse N de modo que se satisfaga el teorema de muestreo, para lo cual debe ser $f_s > 2f_{m\acute{a}x}$, es decir

$$N > \frac{4L}{c} f_{m\acute{a}x}. \quad (78)$$

Por ejemplo, para $f_{m\acute{a}x} = 5000$ Hz y una longitud típica del tracto vocal de $L = 0,17$ m se obtiene un valor $N \geq 10$.

Bajo las condiciones anteriores, llamando $t_n = nT$ y muestreando la ecuación (77) en estos instantes particulares se cumple que

$$\sum_{k=0}^N a_k u_{labios}(nT - kT) = u_{glotis}\left(nT - \frac{N}{2}T\right), \quad (79)$$

y, pasando al tiempo discreto (con abuso de notación),

$$\sum_{k=0}^N a_k u_{labios}(n - k) = u_{glotis}\left(n - \frac{N}{2}\right) \quad (80)$$

Aplicando transformada Z y operando resulta.

$$G_U(z) = \frac{U_{labios}}{U_{glotis}} = \frac{z^{-N/2}}{\sum_{k=0}^N a_k z^{-k}}. \quad (81)$$

El retardo $z^{-N/2}$ representa el tiempo requerido por la onda sonora generada en la glotis para atravesar el tracto vocal. Tanto desde el punto de vista del análisis como desde el punto de vista de la síntesis dicho tiempo en general carece de importancia, y se lo elimina. Asimismo, se suele normalizar a 1 el término independiente del denominador, obteniéndose el *modelo de predicción lineal* (LPM):

$$G_U(z) = \frac{1}{1 + \sum_{k=1}^N a_k z^{-k}}. \quad (82)$$

Obsérvese que si bien la deducción de este modelo fue hecha para el caso en el que no hay pérdidas ni impedancia de radiación, el mismo permite representar también sistemas disipativos.

Bibliografía y referencias

1. Beranek, Leo L. "Acústica". Editorial Hispanoamericana. S. A. Buenos Aires (Arg.), 1961
2. Madisetti, Vijay K.; Williams, Douglas B. (editores). "The Digital Signal Processing Handbook". CRC-Press - IEEE Press. Boca Raton, FL, USA, 1997. Section X.
3. Sondhi, M. M. "Model for wave propagation in a lossy vocal tract". Journal of the Acoustical Society of America. Vol 56, No 5, May 1974, pp. 1070-1075.
4. Syrdal, A.; Bennet, R.; Greenspan, S. "Applied Speech Technology" CRC Press. Boca Raton, FL, USA, 1995
5. Pelton, Gordon E., "Voice Processing". MacGraw-Hill. Singapore, 1992.
6. Fletcher, Harvey, "Speech and Hearing in Communication". American Institute of Physics. Woodbury, USA, 1995.